

Genomic Standards Consortium Projects

Dawn Field^{1,2}, Peter Sterk², Renzo Kottmann³, J. Wim De Smet⁴, Linda Amaral-Zettler⁵, Guy Cochrane⁶, James R. Cole⁷, Neil Davies^{8,9}, Peter Dawyndt⁴, George M. Garrity¹⁰, Jack A. Gilbert^{11,12}, Frank Oliver Glöckner³, Lynette Hirschman¹³, Hans-Peter Klenk¹⁴, Rob Knight¹⁵, Nikos Kyrpides¹⁶, Folker Meyer^{11,17}, Ilene Karsch-Mizrachi¹⁸, Norman Morrison¹⁹, Robert Robbins²⁰, Inigo San Gil²¹, Susanna Sansone², Lynn Schriml²², Tatiana Tatusova¹⁸, Dave Ussery²³, Pelin Yilmaz³, Owen White²², John Wooley²⁰, and Gregory Caporaso²⁴

- ¹ Centre for Ecology and Hydrology, Maclean Building, Benson Lane, Crowmarsh Gifford, Wallingford, Oxfordshire, OX10 8BB, United Kingdom
- ² Oxford e-Research Centre, University of Oxford, 7 Keble Road, Oxford, OX1 3QG, United Kingdom
- ³ Microbial Genomics Group, Max Planck Institute for Marine Microbiology, D-28359 Bremen & Jacobs University Bremen, Germany
- ⁴ Department of Applied Mathematics and Computer Science, Ghent University, Ghent, Belgium
- ⁵ The Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, Woods Hole, Massachusetts, USA
- ⁶ European Molecular Biology Laboratory (EMBL) Outstation, European Bioinformatics Institute (EBI), Wellcome Trust Genome Campus, Hinxton, Cambridge, United Kingdom
- ⁷ Center for Microbial Ecology, Michigan State University, East Lansing, Michigan, USA
- ⁸ Gump South Pacific Research Station, University of California Berkeley, BP 244 98728 Moorea, French Polynesia
- ⁹ Biodiversity Institute, Department of Zoology, University of Oxford, The Tinbergen Building, South Parks Road, Oxford, OX1 3PS, United Kingdom
- ¹⁰ Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, Michigan, USA
- ¹¹ Institute for Genomic and Systems Biology, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, IL 60439, USA
- ¹² Department of Ecology and Evolution, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637, USA
- ¹³ Information Technology Center, The MITRE Corporation, Bedford, Massachusetts, USA
- ¹⁴ DSMZ - German Collection of Microorganisms and Cell Cultures GmbH, Braunschweig, Germany
- ¹⁵ Howard Hughes Medical Institute, and BioFrontiers Institute, and Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado, USA
- ¹⁶ DOE Joint Genome Institute, Walnut Creek, California, USA
- ¹⁷ Computation Institute, University of Chicago, Chicago, Illinois, USA
- ¹⁸ National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, Maryland, USA
- ¹⁹ University of Manchester, Oxford Road, Manchester, United Kingdom
- ²⁰ University of California San Diego, La Jolla, California USA
- ²¹ LTER Network Office, Department of Biology, University of New Mexico, Albuquerque, New Mexico, USA
- ²² Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, MD 20742, USA
- ²³ DOE Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA
- ²⁴ Center for Microbial Genetics and Genomics, Northern Arizona University, Flagstaff, Arizona, USA.

The Genomic Standards Consortium (GSC) is an open-membership community that was founded in 2005 to work towards the development, implementation and harmonization of standards in the field of genomics. Starting with the defined task of establishing a minimal set



of descriptions the GSC has evolved into an active standards-setting body that currently has 18 ongoing projects, with additional projects regularly proposed from within and outside the GSC. Here we describe our recently enacted policy for proposing new activities that are intended to be taken on by the GSC, along with the template for proposing such new activities.

The Genomic Standards Consortium

The Genomic Standards Consortium (GSC) is an open-membership community working towards the development, implementation and harmonization of standards in the field of genomics. The mission of the GSC is to improve digital descriptions of genomes, metagenomes and gene marker sequences. The GSC started in late 2005 with the defined task of establishing what is now termed the “Minimum Information about any Sequence” (MIxS) standard [1,2]. As an outgrowth of the activities surrounding the creation and implementation of the MIxS standard there are now 18 projects within the GSC [3]. These efforts cover an ever widening range of standardization activities. Given the growth of projects and to promote transparency, participation and adoption the GSC has developed a “GSC Project Description Template”. A complete set of GSC Project Descriptions and the template are available on the GSC website. The GSC has an open policy of participation and continues to welcome new efforts. Any projects that facilitate the standard descriptions and exchange of data are potential candidates for inclusion under the GSC umbrella. Areas that expand the scope of the GSC are encouraged. Through these collective activities we hope to help foster the growth of the ‘bioinformatics standards’ community. For more information on the GSC and its range of projects, please see <http://gensc.org/>.

Initiating and Maintaining a Project within the GSC

The GSC Project Description template provides a uniform set of statements covering the purpose of the project, the main contacts, the relationship to other projects in the GSC, and other key details that will help the GSC and the wider community to join and use the work of each project group.

Active projects must update their project descriptions annually and submit them to the GSC Board

References:

1. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, Thomson N, Allen MJ, Angiuoli SV, *et al.* The minimum information

for review. Over time, the list of GSC projects will evolve as activities are added, merged, split or deprecated as demand dictates. Each project should maintain a balanced set of members, especially when the primary goal is to create consensus-driven data sharing solutions.

Project Chairs and members form the core of the GSC. Chairs of projects/working groups report to the GSC Board and help drive GSC workshops and other events. Project working groups define the strategy of the GSC, plan core GSC workshops and host satellite meetings. All working members of GSC projects are encouraged to use GSC events as their primary working forum.

To initiate a new GSC project, or to bring an existing project into the GSC, one or more “champions” prepare a Project Description and present it to the GSC Board. This starts the process of establishing the project within the GSC. Projects are formally voted into the GSC at annual meetings. All proposed GSC activities are open to challenge and improvement over time, especially because the technologies that underpin data sharing change rapidly. Ideally, a balanced working group will be established for all GSC projects, which should be consensus-driven.

By coming together as a community with a common goal and shared interests, we will strengthen and accelerate work towards the goals of establishing best-practices in the capture and exchange of contextual information. The GSC helps grow an ecosystem of standards, tools and resources for empowering a broad community of researchers to describe and share their data. This in turn accelerates the standardization and exchange of an increasing number of data types, and in the long run, promotes open and facile data sharing, thus resolving a current bottleneck for many projects and investigators.

about a genome sequence (MIGS) specification. *Nat Biotechnol* 2008; **26**:541-547.
[Med http://dx.doi.org/10.1038/nbt1360](http://dx.doi.org/10.1038/nbt1360)

2. Yilmaz P, Kottmann R, Field D, Knight R, Cole JR, Amaral-Zettler L, Gilbert JA, Karsch-Mizrachi I, Johnston A, Cochrane G, *et al.* Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. *Nat Biotechnol* 2011; **29**:415-420. [PubMed](#)
3. Field D, Amaral-Zettler L, Cochrane G, Cole JR, Dawyndt P, Garrity GM, Gilbert J, Glockner FO, Hirschman L, Karsch-Mizrachi I, *et al.* The Genomic Standards Consortium. *PLoS Biol* 2011; **9**:e1001088. [PubMed](#) <http://dx.doi.org/10.1038/nbt.1823>
<http://dx.doi.org/10.1371/journal.pbio.1001088>